The effect of pain on task switching: pain reduces accuracy and increases reaction times across multiple switching paradigms

Nina Attridge[a]*, Edmund Keogh[b] and Christopher Eccleston[b]

[a]School of Science, Loughborough University

[b]Centre for Pain Research, University of Bath

* Corresponding author:

Nina Attridge

School of Science

Loughborough University

LE11 3TU

n.f.attridge@lboro.ac.uk

Introduction: 491

Discussion: 1493

Tables: 5

Figures: 2

Abstract

Pain disrupts attention, which may have negative consequences for daily life for people with acute or chronic pain. It has been suggested that switching between tasks may leave us particularly susceptible to pain-related attentional disruption, because we need to disengage our attention from one task before shifting it onto another. Switching tasks typically elicit lower accuracies and/or longer reaction times when participants switch to a new task compared to repeating the same task, and pain may exacerbate this effect. We present three studies to test this hypothesis. In Study 1, participants completed two versions of an alternating runs switching task under pain free and thermal pain induction conditions. Pain did not affect performance on either task. In Studies 2 and 3, we examined seven versions of the switching task using large general population samples, experiencing a variety of naturally-occurring pain conditions, recruited and tested on the internet. On all tasks, participants with pain had longer reaction times on both switch and repeat trials compared to participants without pain, but pain did not increase switch costs. In Studies 2 and 3, we also investigated the effects of type of pain, duration of pain, and analgesics on task performance. We conclude that pain has a small dampening effect on performance overall on switching tasks. This suggests that pain interrupts attention even when participants are engaged in a trial, not only when attention has been disengaged for shifting to a new task set.

Keywords: pain; attention; task switching

1. Introduction

Pain demands attention, thereby disrupting performance on other attention-demanding tasks [11; 13; 21; 29-31; 39]. This can be distressing for patients with pain [1; 6; 17], and may have negative consequences for daily life in people with acute or chronic pain.

One aspect of attention that may be particularly susceptible to disruption from pain is switching between tasks [12; 39]. When we switch between tasks, we must first disengage our attention. While cognitive engagement can act as a distraction from pain [24], this disengagement and shifting may leave attention susceptible to intrusion from pain, thereby increasing the difficulty of switching between tasks.

Eccleston [12] asked chronic pain patients and healthy controls to complete a switching task. Typically, participants have lower accuracies and longer reaction times when switching between tasks compared to repeating the same task. Patients with high-intensity pain were slower to respond to both repeat and switch trials compared to pain-free participants, whereas patients with low-intensity pain did not differ from controls [12]. Pain therefore dampened performance overall, as opposed to increasing the cost of switching [12; 29; 39].

This version of the switching paradigm has since been criticised [28] and alternative versions, such as the task cueing [27] and alternating runs paradigms [35], have become more popular. Moore et al [29] gave participants the task cueing paradigm under three conditions: no pain, warm non-painful heat induction, and painful heat induction. They found that switch costs increased in the warm condition compared to baseline, but there was only a trend towards increased switch costs in the pain condition. On the same task, headache [31] and menstrual pain [21] were found to decrease accuracy on both switch and repeat trials, but again they did not increase the switch cost.

Whilst there are theoretical reasons to hypothesise that pain should increase switch costs, the evidence does not support this. However, each of these studies used a small sample experiencing one type of pain. We aimed to more thoroughly test the hypothesis that pain increases switch costs by employing multiple versions of the switching paradigm in participants experiencing various types of pain. In Study 1, we investigated the effect of pain on two versions of the alternating runs paradigm, which has not previously been investigated in pain. In Study 2, we recruited a large general population sample experiencing various naturally-occurring pain conditions to complete three versions of the switching paradigm online. In Study 3, we increased the task complexity and recruited another large general population sample to complete one of four versions of the switching task online. Given that women experience greater pain, and report more attentional disruption than men [2], we investigated sex differences in attention disruption in each study[10; 15]. We predicted that women would show greater attentional disruption than men. Based on theory [12; 29] as opposed to previous findings [12; 21; 29; 31], we predicted that pain would increase the cost of switching on all versions of the task.

2. Study 1 Method

2.1. Participants

Participants were 44 staff and students from the University of Bath (22 female), aged 18-51 (M=23.75, SD=7.63). All reported being free from pain and not on any medication at the time of the study.

2.2. Design

The experiment followed a 2 (Condition: pain, no pain) x 2 (Task: cued, uncued) x 2 (Sex: male, female) design. Participants completed two versions of an alternating runs

switching task, one with cues and one without. Participants completed both versions of the switching task twice, once while experiencing heat pain and once while pain free. The order of pain condition and tasks was counterbalanced.

2.3. Measures

2.3.1. Switching task overview

In the original version of the task switching paradigm, known as Jersild's method, participants complete blocks of repeat trials (i.e. one block of Task A repeatedly and one block of Task B repeatedly) and blocks of switch trials (continually alternating between Tasks A and B). This was the method used by Eccleston [12] but it has since been criticised because the task demands are unequal between the repeat and switch blocks. In the switch block, participants must keep two instruction and response sets active, while in each repeat block only one instruction and response set is required, so the switch block has a higher arousal and working memory load [28]. In the alternative 'task cueing' paradigm [27], switches occur at random and participants see a cue before each trial telling them which task to perform next, which is sometimes the same as the previous trial and sometimes different. Thus, the task demands are equal for both trial types, except for the key manipulation of switch versus repeat. In the 'alternating runs' paradigm [35], switches occur every *n* trials (typically two), so participants perform the tasks in a set order of AABBAABB, and so forth, either with or without cues before each trial (thus varying the load on working memory). Again, the non-switch task demands are equal between trial types. The difference between the task cueing and alternating runs paradigms is the predictability of when switches will occur.

The effects of pain on switching have not previously been investigated using the alternating runs version of the switching paradigm. We therefore began by investigating the

effects of laboratory-induced pain on task switching using a cued and uncued version of the alternating runs paradigm, before using the task in the large-scale Internet studies reported below.

2.3.2. Cued alternating runs task

Participants performed an alternating runs switching task with a run length of 2, where the two instructional sets were deciding whether a given number was odd or even and deciding whether the number was higher or lower than 5 (from the set 1, 2, 3, 4, 6, 7, 8 and 9). The task was presented using E-Prime Professional 2.0 [36] and responses were collected using a serial response box.

Each trial consisted of a coloured screen (green or blue) for 1000ms followed by the stimuli until response. The stimuli were displayed in the centre of the screen in black size 24 Arial font on a white background with a coloured border (green or blue, to match the previous screen) at the edge of the screen. The colour of the previous inter-trial screen and the border of the stimuli screen indicated which instructional set to perform for that number: green indicated the odd/even instruction and blue indicated the high/low instruction (Figure 1a).

Half of participants began with an odd/even instruction run and half began with a high/low instruction run. The response buttons were the leftmost and rightmost buttons on the serial response box and were counterbalanced across participants.

If a participant lost track of which colour related to which instructional set they could press the middle button on the response box. This directed them to a reminder screen, which stayed on screen until the participant was ready to re-start the task. There were 120 trials and a break was offered every 40 trials. Before starting the experimental trials, participants completed 12 practice trials where written reminders about which instructions to follow

were presented below the stimuli, and 12 practice trials without written reminders. Participants were given the option to repeat the practice block.

2.3.3 Uncued alternating runs task

In the uncued task, each trial consisted of a blank screen for 1000ms, followed by the target number until response. There were no external cues and participants were required to remember the order of instructions and when to switch (Figure 1b). There was an option to press the middle button on the response box if the participant lost track of the instructions. This prompted a screen instructing them to respond from the beginning of the sequence when they continued (i.e. AABB). From the point of re-starting responses were scored according to the new order of instructional sets. There were 120 trials (pressing the middle button did not alter the number of trials given, it only reset the order of responses) and a break was offered every 40 trials. All other aspects of the task and practice trials matched the cued version.

2.3.4 Additional measures

Participants answered demographics questions and completed Visual Analogue Scales on paper for the following questions: “How much pain did you feel during the pain condition of the task?” (anchored by “No pain at all” and “Worst pain imaginable”), “How much pain did you feel when the heat pain was at its most intense?” (anchored by “No pain at all” and “Worst pain imaginable”), and “How intrusive/distracting did the pain seem to you?” (anchored by “Not at all distracting” and “Very distracting”). The VAS scales were included as a manipulation check.

2.4 Apparatus

A Medoc Pathway Advanced Thermal Stimulator (ATS) was used for the purpose of heat pain induction. This equipment is designed for use in clinical and research settings and has built-in safety restrictions. A thermode is placed on the participant's skin and the temperature of the thermode is controlled using the associated software or by a manual trigger controlled by the participant.

The thermode, measuring 30 × 30mm, was strapped to the participant's left ankle between the fibularis longus and extensor digitorum longus muscles, slightly above the lateral malleolus. The baseline temperature of the thermode was set at 32°C. To find the participant's heat pain threshold, they were instructed in how to use a manual trigger, and were asked to increase the temperature until it first felt painful. At this point, the temperature was recorded and the heat stimulus was returned to baseline. This procedure was repeated three times, and the mean temperature at which pain was first reported was taken as the participants' heat pain threshold. This temperature was then used to set individual temperature levels for the pain induction procedure.

A 'pulses' programme was used to induce pain during the pain conditions of the switching tasks. The programme started with a baseline temperature of 32°C and after 1 second produced a pulsating heat for 5 seconds, before returning to baseline. The return to baseline temperature took approximately 1 second, the temperature was held at baseline for 1 second, and the temperature took approximately 1 second to increase back to the pulsing segment. During the pulsing segment, the temperature fluctuated between 1°C below and 1°C above the participant's pain threshold 10 times. This cycle of 5 seconds of pulses with 3 second long dips to baseline was repeated throughout the task. The thermode remained on the ankle at the baseline temperature (32°C) during the no pain conditions. The thermode was moved half way around the ankle to slightly above the medial malleolus for the second

task so as to reduce the risk of side effects from prolonged heat exposure on one area of skin.

2.5 Procedure

The study was approved by the Department of Psychology Ethics Committee and the Research Ethics Approvals Committee for Health at the University of Bath. Participants were fully informed of the study procedure and gave written consent. They were able to withdraw at any time.

After being welcomed to the laboratory and giving informed consent, participants performed the threshold-finding procedure described above. Participants then completed the cued and uncued switching tasks twice, once while pain free and once while experiencing a painful heat stimulus, as described above. After completing the switching tasks participants completed the VAS scales on paper and the demographics questions on the computer using E-Prime Professional 2.0[36]. When all tasks had been completed the participants were paid and debriefed.

2.6 Analysis

Individual trials with RTs of less than 200ms or greater than 2000ms were considered erroneous and removed from the analysis. Furthermore, group data were screened for outliers (mean scores more than three standard deviations away from the group mean).

The main hypothesis, that switch costs would be greater in the pain condition than the non-pain condition on both tasks but to a greater extent on the uncued task, was tested with two 2 (Condition: pain, no pain) x 2 (Trial Type: switch, repeat) x 2 (Task: cued, uncued) x 2 (Sex: male, female) ANOVAs, one on accuracy scores and one on reaction times for correct trials.

3. Results

3.1 Data cleaning

Twenty-two participants reported getting lost on the uncued task between 1 and 9 times each. The median number of times participants became lost did not differ between the no pain (Median = 0) and pain conditions (median = 0), Mann-Whitney $U = 65.5$, $z = -1.26$, $p = .207$. Data were screened for extreme RTs (<200ms or >2000ms) and 7.47% of trials were deleted on this basis (the number of extreme RTs did not differ between the pain and no pain conditions, $t(43) = .43$, $p = .672$). Next, mean accuracies for each participant for each task were compared to chance level using a binomial probability calculator. Eight participants did not score above chance level on the uncued task in both the pain and no pain conditions. One additional participant did not score above chance level in the no pain condition only, and three additional participants did not score above chance level in the pain condition only. These 12 participants' accuracy and RT data were removed from the analyses. A McNemar test showed that the proportion of participants performing above chance level did not significantly differ between the conditions, $p = .625$. No participants scored below chance on the cued task. Finally, scores that were more than three standard deviations from the group mean were removed from the analysis. One participant was removed on the basis of low accuracy scores on the cued task. One participant was removed due to long RTs on the cued task. This left 31 participants in the accuracy analysis and 31 in the RT analysis. Missing cases were excluded pairwise to maximise statistical power.

3.2 Pain thresholds and VAS ratings.

The mean pain threshold was 45.64°C (*SD* = 1.93), and was significantly higher in males (*M* = 46.36°C, *SD* = 1.40) than in females (*M* = 44.91°C, *SD* = 2.14), $t(42) = 2.67$, $p = .011$.

The mean VAS response to the question "How much pain did you feel during the pain condition of the task?" was 46.66 out of 100 (*SD* = 18.58). The mean response to the question "How much pain did you feel when the heat pain was at its most intense?" was 56.16 (*SD* = 20.09). Finally, the mean response to the question "How intrusive/distracting did the pain seem to you?" was 45.00 (*SD* = 25.97). None of the VAS ratings differed by sex (all *p*s >.20)

3.3 The effect of pain on switching

Accuracies and reaction times (shown in Table 1) were entered into separate 2 (Pain) x 2 (Task) x 2 (Trial type) x 2 (Sex) ANOVAs. For accuracies, there was no main effect of Pain, $F(1,29) = 1.08$, $p = .307$, $\eta_\rho^2 = .036$, no main effect of Trial type, $F(1,29) = 1.09$, $p = .305$, $\eta_\rho^2 = .036$, and no main effect of Sex, $F(1,29) = .06$, $p = .814$, $\eta_\rho^2 = .002$, but there was a main effect of Task, $F(1,29) = 12.90$, $p = .001$, $\eta_\rho^2 = .308$, with higher accuracies on the cued task (*M* = .975, *SD* = .016) than on the uncued task (*M* = .926, *SD* = .075). There were no interactions, all *p*s > .130.

For reaction times, there was no main effect of Pain, $F(1,29) = 3.12$, $p = .088$, $\eta_\rho^2 = .097$, no main effect of Task, $F(1,29) < .001$, $p = .998$, $\eta_\rho^2 < .001$, and no main effect of Sex, $F(1,29) = 2.07$, $p = .161$, $\eta_\rho^2 = .067$, but there was a significant main effect of Trial type, $F(1,29) = 22.49$, $p < .001$, $\eta_\rho^2 = .437$, with longer RTs on switch trials (M = 900.64, SD = 192.67) than on repeat trials (M = 790.76, SD = 167.83). There was a significant interaction between Task and Trial type, $F(1,29) = 15.51$, $p < .001$, $\eta_\rho^2 = .348$. This was due to a significant difference between switch (*M* = 946ms, *SD* = 212ms) and repeat (*M* = 758ms,

SD = 182ms) trial RTs on the uncued task, $t(31) = 10.17$, $p < .001$, but not on the cued task, $t(42) = .92$, $p = .365$ (switch *M* = 842ms, *SD* = 219ms; repeat *M* = 815ms, *SD* = 196ms). Finally, there was an interaction between Task, Trial Type and Pain, $F(1,29) = 4.60$, $p = .041$, $\eta_\rho^2 = .037$. However, a series of follow up analyses to identify the source of the interaction showed no two way interactions between trial type and pain within the cued ($p = .377$) or uncued ($p = .694$) task, nor between task and pain within repeat ($p = .305$) or switch trials ($p = .108$), and there were significant interactions between task and trial type in both the pain ($p = .005$) and no pain ($p < .001$) conditions (this interaction was described above; there was a switch cost on the uncued task but not the cued task). Therefore, we were unable to break down the three-way interaction. There were no other interactions, all *p*s > .088.

4. Study 1 discussion

We investigated the effect of pain on performance on two versions of an alternating runs switching task. The tasks involved switching between two instruction sets every two trials. In one version of the task, participants were reminded which task to perform before each trial, whereas in the second version participants did not see reminders but had to remember to keep switching instruction sets every two trials. There was a significant reaction time switch cost on the uncued task, but no switch cost on the cued task. In line with previous research using other versions of the task switching paradigm [12; 21; 29; 31], pain did not increase switch costs on either the cued or uncued version of the task.

To date, research has failed to show a consistent increase in switch costs during pain on four versions of the switching paradigm [12; 21; 29; 31]: Jersild's method, task-cueing, cued alternating runs and uncued alternating runs. However, each of these studies suffered from several limitations: each relied on small samples, including Study 1 here [12; 21; 29; 31], Eccleston's study used Jersild's method [12] which has since been critiqued, and Moore

at al [29] and Study 1 both used laboratory-induced pain which has a low threat value, as opposed to natural-occurring pain which is more threatening, less predictable, and uncontrollable. Study 2, therefore, investigated the effect of pain on the task-cueing, uncued alternating runs, and cued alternating runs tasks in a large (1000+), heterogeneous, general population sample, reporting a variety of pain conditions. Recruiting a large general population sample online has several benefits: the pain is naturally-occurring and therefore has characteristics that are lacking in induced pain, such as threat and uncontrollability, we are able to see the effects of pain on task performance in more naturalistic conditions (i.e. with participants in their everyday environments), and we are able to examine the role of factors such as age, type of pain, and duration of pain, which tend to be homogenous in small samples.

5. Study 2 Method

5.1. Participants

In order to address the issue of small sample sizes, we examined various recruitment methods. One approach that is growing in popularity is to make use of the Internet for recruitment and data collection (in the context of pain, see [2; 3; 8]). Amazon's Mechanical Turk (MTurk) is an open online marketplace for recruiting individuals to complete tasks for a small fee. Research suggests that data collected from MTurk samples is valid and reliable [9; 33]. Users tend to be internally motivated and complete tasks to a high standard for little external reward [26; 33]. MTurk samples also tend to be more demographically diverse than traditional samples in psychology research (i.e. university students and clinical samples) [16; 25]. We therefore decided to use MTurk to recruit participants for Study 2.

A total of 1254 participants accessed the study webpages, 1000 of which were recruited directly through MTurk and 254 of which learned about the study through MTurk

forums and took part outside of the MTurk system. Participants who took part directly through MTurk were paid $2.00; all other participants were unpaid.

Of the 1254 participants who opened the information and consent page, 1088 gave consent to take part, and 1078 of those said that they wanted to seriously participate. These 1078 participants form our sample. There were 554 males, 519 females, 4 male-to-female transsexuals, and 1 female-to-male transsexual. Ages ranged from 18 to 75 (M = 35.76, SD = 11.47). Participants reported their country of residence (USA = 811; India = 230; fewer than 10 participants each from other countries), ethnicity (522 were White American, 235 were Asian Indian, 141 were White European, 42 were Black African, 24 were Other Asian ethnicities, and there were fewer than 20 each from 10 other ethnic groups), and native language (902 English, 58 Hindi, 7 Bengali, and a variety of other languages reported by three or fewer participants each).

5.2. Design

The study followed a mixed groups design, with participants completing all three switching tasks, in a randomised order, and pain status varying between-participants. The study was run using Qualtrics [34] with the switching tasks supported by the QRTEngine [5], to allow for accurate reaction time recording. The QRTEngine has been shown to provide accurate reaction time measurement, and has successfully reproduced classic reaction time effects in a Stroop task, an attentional blink task, and a masked-priming task [25].

5.3. Measures

5.3.1. Seriousness check

Seriousness checks can be used to improve data validity in online research by identifying non-serious submissions [4]. Before completing the experiment participants were asked to indicate whether they wanted to seriously participate or just browse the study pages. At the end of the study, participants were also asked whether they had answered the questions honestly (yes (N = 971), mostly yes (N = 16), mostly no (N = 0), no (N = 1), missing data (N = 90)).

5.3.2. Demographics

Participants were asked to provide demographic information including age, sex, native language, country of residence, ethnicity, and level of education.

5.3.3. Pain status

Participants were asked whether or not they were currently experiencing any pain, and whether they had any recurrent (but not current) pain. This data was used to form three groups: current pain group (participants who were in pain at the time of the study, who may also have had an additional recurrent pain condition that was not causing pain at that moment in time), recurrent pain group (participants who had a recurrent pain condition, but were not in pain at the time of the study), and a no pain group (participants who had neither current nor recurrent pain). If participants answered yes to either of these questions they were shown a series of follow up questions. They were asked to indicate the intensity of their (re)current pain on an 11-point scale labelled 0 ‘no pain at all’ to 10 ‘pain could not be worse’, the type(s) of pain they were experiencing by selecting any applicable items from a list or entering any other conditions in a free text box, the duration of their current pain (up to an hour, up to 24 hours, up to a month, up to three months, up to six months, up to a year, up to a decade, over a decade), whether they had a diagnosis for their pain, whether they

were currently taking analgesics, and how they tend to deal with their pain (by ticking boxes for all applicable strategies: pain killers, distraction, ignore it, go to bed, alcohol/drugs, relaxation/meditation, hot/cold treatments, acupuncture, herbal remedies, exercise).

5.3.4. Experience of Cognitive Intrusion of Pain Scale

Participants completed the Experience of Cognitive Intrusion of Pain (ECIP) scale [2] to measure the phenomenology of cognitive intrusion from pain. However, these data were intended for a separate and ongoing collection of ECIP data, and are not reported here.

5.3.5. Switching tasks

5.3.5.1. Overview of switching tasks

Participants completed three versions of the switching task taken from previous research: the task-cueing paradigm and two versions of the alternating runs paradigm, one with cues and one without. These two paradigms were developed to replace Jersild's paradigm, which suffered from unequal task demands between switch and repeat trials.

5.3.5.2. Cued unpredictable switching task

Participants performed a task-cueing style switching task where the two instructional sets were to decide whether a given number was odd or even and to decide whether it was higher or lower than 5 (from the set 1, 2, 3, 4, 6, 7, 8 and 9). Participants were told which task to perform before each trial (with the cues "Lower or higher than 5?" and "Odd or even?"), and switches occurred pseudo-randomly (according to one of six pre-set trial lists, to which participants were randomly assigned).

Each trial consisted of a fixation cross for 500ms, followed by a task cue for 1000ms, followed by the stimuli until response. The stimuli were displayed in the centre of the screen

in black Arial font on a white background. The number appeared in a 320-pixel × 230-pixel box with a black outline and white background. Participants responded using the ‘c’ (odd or lower than 5) and ‘m’ (even or higher than 5) keys on their keyboard.

There were 64 experimental trials. Before starting the real trials, participants completed 10 practice trials and were given the option to repeat the practice block as many times as they liked.

5.3.5.3. Cued predictable switching task

Participants performed a similar cued alternating runs switching task to that described in Study 1. Again, the two instructional sets were deciding whether a given number was odd or even and deciding whether a number was higher or lower than 5 (from the set 1, 2, 3, 4, 6, 7, 8 and 9). Switches occurred every two trials and participants saw a cue before each trial informing them of which task to perform on that trial. However, in this study the cues took the form of written instructions (“Lower or higher than 5?” and “Odd or even?”) rather than colour codes.

Each trial consisted of a fixation cross for 500ms, followed by a task cue for 1000ms, followed by the stimuli until response. The stimuli were displayed in the centre of the screen in black Arial font on a white background. The number appeared in a 320-pixel × 230-pixel box with a black outline and white background. Participants responded using the ‘c’ (odd or lower than 5) and ‘m’ (even or higher than 5) keys on their keyboard.

There were 64 experimental trials. Before starting the real trials, participants completed 10 practice trials and were given the option to repeat the practice block as many times as they liked.

5.3.5.4. Uncued predictable switching task

Participants performed a similar uncued alternating runs switching task to that described in Study 1, where again the two instructional sets were deciding whether a given number was odd or even and deciding whether a number was higher or lower than 5 (from the set 1, 2, 3, 4, 6, 7, 8 and 9). Switches occurred every two trials and participants did not see cues before each trial, rather they were instructed to remember the sequence of tasks: odd/even, odd/even, low/high, low/high, and so on.

Each trial consisted of a fixation cross for 1500ms (to make the inter-trial-interval equal to the other two tasks, which each had a 500ms fixation cross and a 1000ms cue), followed by the stimuli until response. The stimuli were displayed in the centre of the screen in black Arial font on a white background. The number appeared in a 320-pixel × 230-pixel box with a black outline and white background. Participants responded using the 'c' (odd or lower than 5) and 'm' (even or higher than 5) keys on their keyboard. If participants lost track of the order of tasks, they could press the Q key. This prompted a screen instructing them to respond from the beginning of the sequence when they continued (i.e. odd/even, odd/even, low/high, low/high, and so on.). From the point of re-starting, responses were scored according to the new order of instructional sets.

There were 64 experimental trials (pressing the Q key did not alter the number of trials given, it only reset the order of responses). Before starting the real trials, participants completed two blocks of 10 practice trials. In the first block there were task cues before each trial to reinforce the sequence. In the second block there were no cues. Participants were given the option to repeat the practice blocks as many times as they liked.

5.3.6. Environment

Because participants took part in the study outside of the laboratory, they were asked a series of questions regarding the environment in which they completed the study to allow us

to examine any influence these factors may have on task performance. Participants reported where they completed the study (home (N = 853), work (N = 121), internet café (N = 3), library (N = 10), public transport (N = 0), outdoors (N = 0), other (N = 4), missing data (N = 87)), whether they were interrupted (no (N = 825), once (N = 121), a few times (N = 39), repeatedly (N = 4), missing data (N = 89)), the amount of noise in their environment (on a scale of 0 (silent) to 10 (very noisy), M = 2.30, SD = 1.92, missing data N = 89), and how much effort they put into the tasks (also on a scale of 0 (no effort) to 10 (as much effort as possible), M = 9.49, SD = .90, missing data N = 87). We present an analysis of how these factors affected task performance in the results section.

5.4. Procedure

The study was approved by the University of Bath Departments of Psychology and Health ethics boards. Participants were recruited via Amazon's Mechanical Turk, pain discussion forums, or other online advertisements. The study was presented to participants using Qualtrics with the QRTEngine extension to allow for reaction time (RT) recording [5]. The study opened with an information and consent page. Participants were required to check a box and click next if they consented to take part. If they did not consent to take part, they were directed to an exit page and did not see any of the study materials. Participants who gave consent completed the following sections in order: demographics, pain questions, ECIP scale, the three switching tasks (in a randomised order) and the questions about their environment. Finally, participants saw a debrief page and received a completion code for the MTurk payment system. The experiment took approximately 20-25 minutes to complete. All responses were anonymous and participants were able to withdraw at any time.

5.5. Analysis

Sex differences in the prevalence of current and recurrent pain were investigated using Chi-Square tests. Within the participants who reported pain, sex differences in intensity were investigated with t-tests. The relationship between age and pain was investigated with a logistic regression analysis.

The relationship between participants' environment and their task accuracies and RTs was investigated using regression models with five independent variables: noise, interruptions, location (e.g. home, work), honesty and effort.

The effect of pain on accuracy and RT scores on the switching tasks was analysed with mixed ANCOVAs with two within subjects factors, Task (random, cued, uncued) and Trial Type (switch, repeat), two between subjects factors, Pain (current pain, recurrent but not current pain, and no pain) and Sex (female, male), and two covariates, Age and Noise (based on the results of the regression of environmental factors onto task performance reported below). For investigations of sex differences, we limit our sample to those who reported being male or female in order to preserve cell sizes.

The relationship between current pain intensity and task performance was investigated by correlating pain intensity with accuracies and RTs on each of the three tasks (see supplementary material).

We investigated the effect of type of pain on task performance in the pain type groups that contained at least 10 participants with data on all three tasks (non-muscular back pain, $N = 18$, headache, $N = 16$, joint pain, $N = 14$ and multiple pain types, $N = 173$) using 4 (pain group) × 3 (task: cued unpredictable, cued predictable, uncued predictable) × 2 (trial type: switch, repeat) ANOVAs for accuracy and RT scores (see supplementary material).

We examined the relationship between current pain duration and task performance using Spearman's rank correlations (see supplementary material).

Finally, we investigated the effects of having recently taken analgesics at the time of the study on accuracies and RTs using 3 (group: no pain and no analgesics, pain and no analgesics, and pain and analgesics) × 3 (task: cued unpredictable, cued predictable, uncued predictable) × 2 (trial type: switch, repeat) × 2 (Sex) mixed ANOVAs (see supplementary material).

6. Results

6.1. Pain characteristics

Participants reported whether they had current pain, whether they had recurrent (but not occurring at that moment) pain, and they reported the intensity of any pain they were experiencing on a 0-10 scale. Four hundred and ninety-seven participants reported no pain, 173 reported recurrent but not current pain, 78 reported current but not recurrent pain, and 337 reported both current and recurrent pain. The latter two groups were combined to form a single 'current pain group' in the analyses for two reasons: firstly, we were primarily interested in the effect of current pain on task performance, and these two groups both experienced current pain, and secondly, only a small number of participants reported current but not recurrent pain. Participants who reported current pain reported an average intensity of 4.88 (*SD* = 1.95), and those who reported recurrent pain reported an average intensity of 5.40 (*SD* = 1.82).

Participants also reported the duration and type of their current and recurrent pain (see Table 2).

Sex differences in the prevalence of current pain were investigated using a Chi-Square test, which revealed that, as expected [18], more women (44%) than men (33%) reported that they were in pain, $\chi^2(1) = 11.39$, $p = .001$. This was also the case for recurrent pain (reported in 52% of women and 42% of men), $\chi^2(1) = 9.30$, $p = .003$.

Within the participants who reported pain, there was no sex difference for the intensity of current pain reported, $t(410) = .26$, $p = .797$ (males: $M = 4.84$, $SD = 1.98$, females: $M = 4.89$, $SD = 1.93$). However, women reported higher intensity recurrent pain ($M = 5.60$, $SD = 1.78$) than men ($M = 5.18$, $SD = 1.83$), $t(493) = 2.59$, $p = .010$.

A logistic regression showed that current pain was slightly, although significantly, more common in older participants than in younger participants, Wald = 4.28, Exp(B) = 1.01, $p = .039$. This was also the case for recurrent pain, Wald = 33.48, Exp(B) = 1.03, $p < .001$.

6.2. Data cleaning

Of the 1254 participants who started the study, 1034 completed the cued random switching task, 1047 completed the uncued alternating runs switching task and 1034 completed the cued alternating runs switching task.

The RT data were screened for extreme values. For the uncued alternating runs switching task, 1201 (1.79%) responses with RTs smaller than 200ms were removed and 1900 (2.84%) responses with RTs longer than 3000ms were removed. A Kruskal-Wallis test showed that there was no difference between pain groups in the median number of extreme RTs (no pain median = 0, recurrent but not current pain median = 0, current pain median = 1), $H(1042) = 3.98$, $p = .264$. For the cued alternating runs switching task, 1257 (1.89%) responses with RTs smaller than 200ms were removed and 1445 (2.18%) responses with RTs longer than 3000ms were removed. There was a higher median number of extreme RTs in the current pain group (median = 2) than in the recurrent or no pain groups (medians = 1), $H(1032) = 10.29$, $p = .016$. For the cued random switching task, 1125 (1.69%) responses with RTs smaller than 200ms were removed and 1033 (1.55%) responses with RTs longer than 3000ms were removed. There was no difference between pain groups in the median

number of extreme RTs (median = 0 in all three groups), $H(1029) = 1.423$, $p = .491$. Accuracies associated with extreme RTs were also excluded.

Mean RTs and accuracies were calculated for each participant who had 20 or more remaining values per trial type after extreme RT values were removed. Participants who had fewer than 20 values per trial type were removed from the analysis (uncued = 41, random = 34, cued = 38). Mean RTs were then examined for outliers. No participants had RTs more than 3 SDs below the mean on any of the tasks. Four participants on the uncued alternating runs task, 6 on the cued alternating runs task and 4 on the cued random switching task had RTs more than 3 SDs above the mean, and were removed from the RT analyses to preserve normality. However, these participants' accuracy data was retained. As such, the sample sizes for reaction time analyses was slightly smaller than for accuracy analyses.

A binomial probability calculator was used to calculate the minimum number of correct trials for performance to be above chance level. Participants who did not score above chance level (40/64 or 62.5% of trials correct) were removed from both the accuracy and RT analyses. On the cued random switching task, 17 participants failed to score significantly above chance level. On the uncued alternating runs task, 170 participants reported losing track of the task at least once (maximum 6 times), but the median number of times that participants lost track did not differ between the three groups, $H(1042) = .18$, $p = .912$. Trials on which participants indicated that they had lost track were removed from the analysis and their remaining data were included. Despite only 170 participants reporting getting lost on the task, 333 participants still did not perform above chance level, which may indicate that participants were not always aware of when they had lost track of the sequence of task switches. On the cued alternating runs task, 22 participants did not perform above chance level. A series of Chi-square tests suggested that above chance performance was not related to current pain status, all ps > .235.

6.3. Environmental effects on task performance

The relationship between the remaining participants' environment and their task performance was investigated using three regression models with five independent variables: noise, interruptions, location (e.g. home, work), honesty and effort, one model to predict accuracy on each of the three tasks. For the uncued task, the model was significant, $R^2 = .030$, $F(5,634) = 3.87$, $p = .002$, and noise, $\beta = -.10$, $p = .041$, and location, $\beta = .12$, $p = .002$, were significant predictors. For the cued task, the model was significant, $R^2 = .030$, $F(5,941) = 5.80$, $p < .001$, and noise was a significant predictor, $\beta = -.11$, $p = .003$. For the random task, the model was significant, $R^2 = .034$, $F(5,947) = 6.69$, $p < .001$, and effort, $\beta = .08$, $p = .021$, and noise, $\beta = -.14$, $p < .001$, were significant predictors. Since noise consistently affected task performance, we included it as a covariate in the analyses below.

6.4. Effects of pain on task accuracies

Accuracy scores (shown in Table 3) were analysed with mixed ANCOVAs with two within groups factors, Task (random, cued, uncued) and Trial Type (switch, repeat), two between groups factors, Pain (current pain, recurrent but not current pain, and no pain) and Sex (female, male), and two covariates, Age and Noise. Due to Task being a repeated measures factor, only participants who had scores remaining on all three tasks, after the exclusion criteria described above were applied, are included in the analysis ($N = 611$; males = 316, females = 295; current pain = 246, recurrent but not current pain = 89, no pain = 276).

For accuracy scores there was a main effect of Task, $F(2,1206) = 33.56$, $p < .001$, $\eta_p^2 = .053$, where scores on the uncued task ($M = .834$, $SD = .148$) were significantly lower than scores on the random ($M = .935$, $SD = .074$) or cued tasks ($M = .933$, $SD = .074$), both ps <

.001, but the cued and random tasks did not differ from each other, $p = .279$. There was a main effect of trial type, $F(1,603) = 9.29$, $p = .002$, $\eta_p^2 = .015$, with higher accuracy on repeat trials ($M = .911$, $SD = .074$) than on switch trials ($M = .891$, $SD = .074$). There was a significant main effect of Age, $F(1,603) = 4.18$, $p = .041$, $\eta_p^2 = .007$, with accuracy increasing as age increased, $r(615) = .120$, $p = .003$, and a main effect of noise, $F(1,603) = 17.86$, $p < .001$, $\eta_p^2 = .029$, with accuracy decreasing as noise increased, $r(611) = -.191$, $p < .001$. Importantly, there was a main effect of Pain, $F(1,603) = 4.74$, $p = .009$, $\eta_p^2 = .015$, with lower accuracy in the pain group ($M = .889$, $SD = .063$) than in the recurrent but not current pain group ($M = .912$, $SD = .066$), p = .004, and the no pain group ($M = .901$, $SD = .063$), $p = .039$. The recurrent but not current pain group did not differ from the no pain group, $p = .159$.

For accuracy scores there was also an interaction between Task and Sex, $F(2,1206) = 4.33$, $p = .013$, $\eta_p^2 = .007$. Independent samples t-tests showed that females ($M = .842$, $SD = .120$) scored higher than males ($M = .821$, $SD = .123$) on the uncued task, $t(613) = 2.12$, $p = .035$, $d = 0.17$, but that males and females did not differ on the random, $t(613) = 1.07$, $p = .287$, d = .09, or cued tasks, $t(613) = 1.04$, $p = .299$, $d = .08$.

Finally, there was a three way interaction between Task, Sex and Pain, $F(2,1206) = 3.10$, $p = .015$, $\eta_p^2 = .010$. However, when we broke this down into two-way interactions, we were unable to locate the source of the interaction: there was no interaction between Task and Pain for males or females, both ps > .630, no interaction between Task and Sex in any of the pain groups, all ps > .120, and no interaction between Sex and Pain on any of the tasks, all ps > .580.

6.5. Effects of pain on task reaction times

Correct RTs (shown in Table 3) were analysed with mixed ANCOVAs with two within groups factors, Task (random, cued, uncued) and Trial Type (switch, repeat), two between groups factors, Pain (current pain, recurrent but not current pain, no pain) and Sex (female, male), and two covariates, Age and Noise. Due to Task being a repeated measures factor, only participants who had scores remaining on all three tasks, after the exclusion criteria described above were applied, are included in the analysis (N = 607 males = 313, females = 294; current pain = 244, recurrent but not current pain = 89, no pain = 274).

For RTs there was a main effect of Trial Type, $F(1,599) = 54.52$, $p < .001$, $\eta_p^2 = .083$, with longer RTs on switch trials (M = 992ms, SD = 241ms) than on repeat trials (M = 890ms, SD = 214ms). There was a main effect of noise, $F(1,599) = 5.21$, $p = .023$, $\eta_p^2 = .009$, with more noise related to longer RTs, $r(607) = .106$, $p = .009$. There was also a main effect of pain, $F(1,599) = 5.99$, $p = .003$, $\eta_p^2 = .020$, with longer RTs in the current pain group (M = 968ms, SD = 197ms) than in the no pain group (M = 908ms, SD = 198ms), $p = .001$. The recurrent but not current pain group (M = 948, SD = 196) did not significantly differ from either the current pain, $p = .413$, or the no pain group, $p = .097$.

For RTs there was also a series of two-way interactions. Task interacted with Trial Type, $F(2,1198) = 5.06$, $p = .007$, $\eta_p^2 = .008$. This was due to a larger difference between repeat and switch trials on the uncued task, $t(611) = 19.17$, $p < .001$, $d = .50$, and the cued task, $t(614) = 22.23$, $p < .001$, $d = .47$, than on the random task, $t(613) = 14.01$, $p < .001$, $d = .26$. Age interacted with Task, $F(2,1198) = 7.23$, $p = .001$, $\eta_p^2 = .012$: age predicted RTs on the uncued task, $r(611) = .108$, $p = .008$, but not on the random, $r(611) = -.016$, $p = .691$, or cued tasks, $r(611) = .004$, $p = .926$. Noise interacted with Trial Type, $F(1,599) = 8.88$, $p = .003$, $\eta_p^2 = .015$: noise predicted RTs on repeat trials, $r(611) = .130$, $p = .001$, but not on switch trials, $r(611) = .064$, $p = .113$.

7. Study 2 discussion

Participants who reported being in pain at the time of the study had lower accuracies and longer reaction times than participants who reported no pain. Participants with recurrent but not current pain did not differ from participants without pain on either accuracies or RTs, but had higher accuracies than participants with current pain. Again, we did not observe an increase in switch costs on any of the tasks, similar to previous studies [12; 21; 29; 31]. There were no two-way interactions between Sex and Pain for accuracies or RTs, so although women may be more susceptible to pain and may report greater experiences of cognitive intrusion from pain [2], we did not find any behavioural evidence that their cognition is more susceptible to the disruptive effects of pain.

In Studies 1 and 2 we failed to find any increased switch costs associated with pain. However, in Study 2 we did find that accuracies and RTs are generally worse in participants with pain. This was not the case in Study 1, although an examination of the means in Tables 1 and 3 show that there was a slight trend in Study 1 in the same direction as the findings for Study 2, and the effect sizes for the equivalent effects were larger in Study 1. Study 2 benefitted from a very large sample, providing sufficient power to detect even small effects.

In Study 3, we attempted to elicit an effect of pain on switch costs by increasing the complexity of the tasks, again using a large sample recruited online. Participants were randomly assigned to complete one of four versions of a switching task using more complex stimuli, similar to that used by Eccleston et al. In each task, participants saw two cards on each trial. Each card had between 1 and 9 digits, and the digits had a value of 1 to 9 (see Figure 2). The two instruction sets were to identify the card with the highest value digits and to identify the card with the largest quantity of digits. The four tasks varied on two dimensions: predictability of switches and interference of instruction sets. By increasing the complexity of the stimuli and varying the difficulty of the tasks, we aimed to elicit an

increase in switch costs in participants with pain compared to those without. In two tasks the switches occurred at random (task-cueing paradigm) and in two they occurred every two trials (alternating runs paradigm). One of each task type required participants to respond by choosing the left or right hand card (low response interference) and one of each task required the participant to return the relevant number (e.g. if the left card has two sevens and the right card has five fours and it was a value trial, the correct response would be the 7 key, whereas for a quantity trial the correct response would be the 5 key; high response interference).

7. Study 3 Method

7.1. Participants

Participants were again recruited through Amazon's Mechanical Turk (MTurk), an open online marketplace for recruiting individuals to complete tasks for a small fee. A total of 4,189 participants opened the study webpages, 3200 of which were recruited directly through MTurk and 989 of which learned about the study through MTurk forums and took part outside of the MTurk system. Participants who took part directly through MTurk were paid $1.50; all other participants were unpaid.

Of the 4189 participants who opened the information and consent page, 3209 gave consent and said that they wanted to seriously participate. One participant reported their age as 15 and was excluded. The remaining 3208 participants form our sample. There were 1603 males, 1544 females, 16 transsexuals and 10 who did not report their sex. Ages ranged from 18 to 75 (M = 35.53, SD = 11.29). Participants reported their country of residence (USA = 2522; India = 608; fewer than 10 participants each from other countries), ethnicity (2052 were White, 765 were Asian, 344 were Black, 110 had mixed ethnicity, and 68 reported

other or unknown), and native language (2731 English, 145 Hindi, 18 Spanish, 10 Bengali, and a variety of other languages reported by fewer than 10 participants each).

7.2. Design

The study was run using Qualtrics [34] with the switching tasks supported by the QRTEngine [5], to allow for accurate reaction time recording. Participants were randomly assigned to complete one of four versions of a switching task. On opening the study link, Google Analytics code randomly redirected each participant to one of four versions of the study. This occurred prior to the participant answering any questions and so the assignment to tasks did not take into account pain characteristics. This led to similar but not equal group sizes for each task. We compared performance between participants with current pain, recurrent but not current pain, and no pain (the tasks were assigned between-participants as opposed to within-participants to keep the study as short as possible to prevent drop-outs).

7.3. Measures

7.3.1. Seriousness check

As in Study 2 participants were asked to indicate whether they wanted to seriously participate or just browse the study pages and at the end of the study they were asked whether they had answered the questions honestly (yes (N = 2837), mostly yes (N = 95), mostly no (N = 7), no (N = 0), missing data (N = 198)).

7.3.2. Demographics, pain status and Experience of Cognitive Intrusion of Pain

Participants answered the same demographic and pain related questions as in Study 2 (sections 5.3.2. and 5.3.3), and they again completed the Experience of Cognitive Intrusion scale (section 5.3.4), which contributed to an ongoing collection of data on this new scale

which is not reported here. The pain groups were formed in the same way as in Study 2: current pain group (participants who were in pain at the time of the study, who may also have had an additional recurrent pain condition that was not causing pain at that moment in time), recurrent pain group (participants who had a recurrent pain condition, but were not in pain at the time of the study), and a no pain group (participants who had neither current nor recurrent pain).

7.3.3. Switching tasks

7.3.3.1. Overview of tasks

In each task, participants saw two cards on each trial. Each card had between 1 and 9 digits, and the digits had a value of 1 to 9 (see Figure 2). The two instruction sets were to identify the card with the highest value digits and to identify the card with the largest quantity of digits. The four tasks varied on two dimensions: predictability of switches (as in Study 2) and interference of instruction sets. In two tasks the switches occurred at random (task-cueing paradigm) and in two they occurred every two trials (alternating runs paradigm). One of each task type required participants to respond by choosing the left or right hand card (low response interference) and one of each task required the participant to return the relevant number (e.g. if the left card has two sevens and the right card has five fours and it was a value trial, the correct response would be the 7 key, whereas for a quantity trial the correct response would be the 5 key; high response interference). We hypothesised that the high response interference condition would make task switches more difficult, and that this additional difficulty would lead to an increase in switch costs in participants with pain compared to those without.

7.3.3.2. Task-cueing paradigm with binary response (“RandBinary” task)

Participants performed a task-cueing style switching task where the stimuli consisted of two cards, each displaying 1 to 9 digits with values of 1 to 9. All digits on a card had the same value, and the two cards in each pair never had the same value digits or the same quantity of digits as each other. The two tasks were to indicate which task had the highest value digits (the "Value" task) and to indicate which task has the greatest quantity of digits (the "Quantity" task). Participants were told which task to perform before each trial (with the cues "Value" and "Quantity"), and switches occurred pseudo-randomly (according to one of three pre-set trial lists, to which participants were randomly assigned).

Each trial consisted of a fixation cross for 500ms, followed by a task cue for 1000ms, followed by the stimuli until response. The stimuli consisted of black digits in Arial font, with a black playing card shaped border. Each card was 320 pixels in height and 230 pixels in width. Participants responded using the 'c' (left) and 'm' (right) keys on their keyboard.

There were 160 experimental trials. Before starting the real trials, participants completed 10 practice trials and were given the option to repeat the practice block as many times as they liked.

7.3.3.3.Task-cueing paradigm with numerical response ("RandNum" task)

The task was identical to the RandBinary task described above, except that participants were required to return a numerical value instead of a binary left/right response. For example, on a Value trial where the left card displayed three fives and the right card displayed seven twos, the correct response would be "5" (i.e. the highest value shown).

7.3.3.4. Alternating runs paradigm with binary response ("AltBinary" task)

The task was identical to the RandBinary task described above, except that the task switches occurred every two trials and participants did not see cues before each trial.

Instead, participants were instructed to remember the sequence of tasks: Value, Value, Quantity, Quantity, etc. If participants lost track of the order of tasks, they could press the Q key. This prompted a screen instructing them to respond from the beginning of the sequence when they continued (i.e. value, value, quantity, quantity). From the point of re-starting responses were scored according to the new order of instructional sets. Pressing the Q key did not alter the number of trials given, it only reset the order of responses. Given that participants did not see cues, the fixation cross was set to 1500ms, to equalise the inter-trial-interval between tasks. The task was otherwise identical to the RandBinary task.

7.3.3.5. Alternating runs paradigm with numerical response ("AltNum" task)

The task was identical to the AltBinary task described above, except that participants were required to return a numerical value instead of a binary left/right response. For example, on a Value trial where the left card displayed six nines and the right card displayed one five, the correct response would be "9" (i.e. the highest value shown).

7.3.4. Environment

Participants were asked a series of questions regarding the environment in which they completed the study: where they completed it (home (N = 588), work (N = 55), internet café (N = 2), library (N = 2), public transport (N = 0), outdoors (N = 2), other (N = 4), missing data (N = 2494)), whether they were interrupted (no (N = 2457), once (N = 382), a few times (N = 100), repeatedly (N = 7), missing data (N = 201)), the amount of noise in their environment (on a scale of 0 (silent) to 10 (very noisy), M = 2.20, SD = 1.66, missing data N = 203), and how much effort they put into the tasks (also on a scale of 0 (no effort) to 10 (as much effort as possible), M = 9.32, SD = 1.18, missing data N = 203). We present an analysis of how these factors affected task performance in the results section.

7.4. Procedure

The study was approved by the University of Bath Departments of Psychology and Health ethics boards. The procedure was identical to Study 2, expect that participants completed one of the four switching tasks, not three tasks. The experiment took approximately 15 minutes to complete. All responses were anonymous and participants were able to withdraw at any time.

7.5. Analysis

Sex differences in the prevalence of current and recurrent pain were investigated using Chi-Square tests. Within the participants who reported pain, sex differences in intensity were investigated with t-tests. The relationship between age and pain was investigated with a logistic regression analysis.

The relationship between participants' environment and their task performance was investigated using regression models with five independent variables: noise, interruptions, location (e.g. home, work), honesty and effort.

The effect of pain on accuracy and RT scores on the switching tasks was analysed with mixed ANCOVAs with two within subjects factors, Task (RandBinary, RandNum, AltBinary, AltNum) and Trial Type (switch, repeat), two between participants factors, Pain (current pain, recurrent but not current pain, no pain) and Sex (female, male), and one covariate, effort (based on the results of the regression of environmental factors onto task performance). For investigations of sex differences, we limit our sample to those who reported being male or female in order to preserve cell sizes.

The relationship between current pain intensity and task performance was investigated by correlating pain intensity with accuracies and RTs on each of the four tasks (see supplementary material).

We investigated the effect of type of pain on task performance in the pain type groups that contained at least 10 participants (which varied by task) using ANOVAs for accuracy and RT scores (see supplementary material).

We examined the relationship between current pain duration and task performance using Spearman's rank correlations (see supplementary material).

Finally, we investigated the effects of analgesics on accuracies and RTs using 3 (group: no pain and no analgesics, pain and no analgesics, and pain and analgesics) × 4 (task: RandBinary, RandNum, AltBinary, AltNum) × 2 (trial type: switch, repeat) × 2 (Sex) mixed ANOVAs (see supplementary material).

8. Results

8.1. Pain characteristics

Participants reported whether they had current pain, whether they had recurrent (but not current) pain, and they reported the intensity of any pain they were experiencing on a 0-10 scale. 1451 participants reported no pain, 534 reported recurrent but not current pain, 214 reported current but not recurrent pain, and 1009 reported both current and recurrent pain. Participants who reported current pain reported an average intensity of 4.95 (*SD* = 2.01), and those who reported recurrent pain reported an average intensity of 5.34 (*SD* = 1.87). Again, participants were assigned to one of three groups: current pain, recurrent but not current pain, and no pain.

Participants reported the duration and type of their current and recurrent pain (see Table 4). Sex differences in the prevalence of current pain were investigated using a Chi-

Square test, which revealed that a higher proportion of women (43%) than men (34%) reported that they were in pain at the time of the study, $\chi^2(1) = 26.60$, $p < .001$. This was also the case for recurrent pain conditions that were not currently painful (reported in 54% of women and 42% of men), $\chi^2(1) = 45.16$, $p < .001$.

Within the participants who reported current pain, there was no sex difference in pain intensity, $t(1197) = .55$, $p = .580$ (males: $M = 5.00$, $SD = 2.07$, females: $M = 4.94$, $SD = 1.95$). However, in the participants who reported recurrent pain, women reported higher intensity ($M = 5.50$, $SD = 1.85$) than men ($M = 5.15$, $SD = 1.87$), $t(1472) = 3.68$, $p < .001$.

A logistic regression showed that current pain was slightly, although significantly, more common in older participants than in younger participants, Wald = 48.24, Exp(B) = 1.02, $p < .001$. This was also the case for recurrent pain, Wald = 95.88, Exp(B) = 1.03, $p < .001$.

8.2. Data cleaning

The RT data were screened for extreme values. For the AltBinary task, 1411 (1.31%) responses with RTs smaller than 200ms were removed and 5744 (5.3%) responses with RTs longer than 3000ms were removed. There were no differences in the prevalence of extreme RTs between pain groups, $p = .527$. For the AltNum task, 1339 (1.1%) responses with RTs smaller than 200ms were removed and 15627 (12.4%) responses with RTs longer than 3000ms were removed. The current pain group had a higher median number of extreme RTs (median = 3) than the recurrent pain or no pain groups (both medians = 2), $H(862) = 13.27$, $p = .001$. For the RandBinary task, 1615 (1.2%) responses with RTs smaller than 200ms were removed and 5297 (3.8%) responses with RTs longer than 3000ms were removed. There were no differences in the prevalence of extreme RTs between pain groups, $p = .783$. For the RandNum task, 1496 (1.2%) responses with RTs smaller than 200ms were removed

and 12721 (9.8%) responses with RTs longer than 3000ms were removed. There were no differences in the prevalence of extreme RTs between pain groups, $p = .453$. Accuracies associated with extreme RTs were also excluded.

On the AltBinary task, 205 participants reported losing track of the task at least once, but a Kruskal-Wallis test suggested that the rate did not differ between the pain and no pain groups, $H(685) = 4.74$, $p = .094$. On the AltNum task 258 participants reported losing track but again, the rate did not differ between the pain and no pain groups, $H(827) = 4.03$, $p = .133$. Trials on which participants indicated that they had lost track were removed from the analysis and their remaining data was included. The fact participants reported getting lost in Study 1 (22/44) as well as in Studies 2 and 3 suggests that the lack of control over participants' environments in the online studies was not a key factor in participants losing track of the order of tasks. Pain also did not seem to be an important predictor of participants losing track. Instead, the high load that the uncued tasks placed on working memory may be responsible.

Mean RTs and accuracies were calculated for each participant who had 20 or more remaining values per trial type after extreme RT values and trials on which they reported being lost were removed. Participants who had fewer than 20 values per trial type were excluded from the analysis (Accuracies: AltBinary = 6, AltNum = 29, RandBinary = 5, RandNum = 19, RTs: Accuracies: AltBinary = 15, AltNum = 200, RandBinary = 7, RandNum = 122).

A binomial probability calculator was used to calculate the minimum number of correct trials for performance to be considered above chance level. On the binary response tasks, there are two response options so the probability of success by chance on each trial is 0.5, meaning 91 (or 56.9%) of the 160 trials must be answered correctly to be considered significantly above chance level. On the numerical response task, there are 9 response

options (the numbers 1 - 9) so the probability of success by chance on each trial is .11. To be considered significantly above chance level, 25 (or 15.6%) of the 160 trials must be answered correctly. Participants who did not score above chance level were removed from both the accuracy and RT the analyses (AltBinary = 13, AltNum = 43, RandBinary = 14, RandNum = 39). A Chi-square test suggested that above chance performance was not related to current pain status, $p = .290$.

8.3. Environmental effects on task performance

The relationship between the remaining participants' environment and their task performance was investigated using a regression model with five independent variables: noise, interruptions, location (e.g. home, work) honesty and effort to predict task accuracy. The model was significant, $R^2 = .027$, $F(5,654) = 3.66$, $p = .002$, and effort was a significant predictor, $\beta = .102$, $p = .014$. Since effort affected task performance, we included it as a covariate in the analyses below.

8.4. Effects of pain on task accuracy

Accuracy scores (shown in Table 5) were analysed with a mixed ANCOVA with one within groups factor, Trial Type (switch, repeat), three between group factors, Task (AltBinary, AltNum, RandBinary, RandNum), Pain (current pain, recurrent but not current pain, and no pain) and Sex (female, male), and one covariate, Effort.

There was a significant main effect of Task, $F(3,2825) = 483.24$, $p < .001$, $\eta_p^2 = .339$, which was due to higher accuracy on the RandBinary task ($M = .917$, $SD = .20$), than on the AltBinary task ($M = .861$, $SD = .20$), which in turn was higher than the RandNum task ($M = .655$, $SD = .22$), which in turn was higher than the AltNum task ($M = .556$, $SD = .21$), all ps $< .001$. There was also a significant main effect of Trial Type, $F(1,2825) =$

11.18, $p = .001$, $\eta_p^2 = .004$, with higher accuracy on repeat trials ($M = .752$, $SD = .21$) than on switch trials ($M = .742$, $SD = .21$). There was a main effect of Sex, $F(1,2825) = 5.62$, $p = .018$, $\eta_p^2 = .002$, which was due to higher accuracy in females ($M = .757$, $SD = .19$) than in males ($M = .738$, $SD = .23$). The main effect of Pain was not significant, $F(1,2825) = 2.87$, $p = .057$, $\eta_p^2 = .002$.

There was a two-way interaction between Trial Type and Effort, $F(1,2825) = 6.90$, $p = .009$, $\eta_p^2 = .002$, where Effort was more strongly related to accuracy on switch trials, $r(2850) = .123$, $p < .001$, than on repeat trials, $r(2850) = .108$, $p < .001$. Finally, there was a four-way interaction between Task, Trial type, Sex and Pain, $F(6,2825) = 2.144$, $p = .046$, $\eta_p^2 = .005$. Breaking this down by Sex revealed a significant interaction between Task, Trial Type and Pain in women, $F(6,1376) = 2.55$, $p = .018$, $\eta_p^2 = .011$, but not men, $F(6,1448) = .730$, $p = .625$. In women, there was a significant interaction between Pain and Trial Type on the AltBinary task, $F(2,299) = 3.22$, $p = .042$, $\eta_p^2 = .021$, but not on any other task, all ps > .180. This was due to a significant difference between accuracies for repeat and switch trials in participants without pain (repeat $M = .855$, $SD = .130$, switch $M = .873$, $SD = .115$), $t(128) = 2.21$, $p = .029$, but no difference for participants with recurrent but not current pain, $t(52) = 1.83$, $p = .072$, or those with current pain, $t(129) = .32$, $p = .752$.

All other interactions were non-significant, all ps > .060.

8.5. Effects of pain on task reaction times

Correct RTs (shown in Table 5) were analysed with a mixed ANCOVA with one within-groups factor, Trial Type (switch, repeat), three between groups factors, Task (AltBinary, AltNum, RandBinary, RandNum), Pain (current pain, recurrent but not current pain, and no pain) and Sex (female, male), and one covariate, Effort.

There was a significant main effect of Task, $F(3,2665) = 645.98$, $p < .001$, $\eta_p^2 = .421$, with the longest RTs on the AltNum task (M = 1605ms, SD = 306ms), followed by the RandNum task (M = 1522ms, SD = 306ms), followed by the AltBinary task (M = 1111ms, SD = 309ms), and finally the RandBinary task (M = 1007ms, SD = 299ms), and all tasks were significantly different from each other, ps < .001. There was a significant main effect of Trial Type, $F(1,2665) = 13.40$, $p < .001$, $\eta_p^2 = .005$, with longer RTs on switch trials (M = 1343ms, SD = 307ms) than on repeat trials (M = 1280ms, SD = 320ms). Finally there was a main effect of Pain, $F(1,2665) = 7.16$, $p < .001$, $\eta_p^2 = .005$, with longer RTs in participants with current pain (M = 1340ms, SD = 279ms) than in participants with recurrent but not current pain (M = 1296, SD = 280), $p = .007$, or those without pain (M = 1298ms, SD = 279ms), $p < .001$. Participants with recurrent but not current pain did not differ from those without pain, p = .909. There was no main effect of Sex, $F(1,2665) = 1.66$, $p = .197$, $\eta_p^2 = .001$.

There was a significant interaction between Task and Trial Type, $F(2,2665) = 50.29$, $p < .001$, $\eta_p^2 = .054$, but RTs were significantly longer for switch trials than for repeat trials on all four tasks, all ps < .001. There was also a significant interaction between Pain and Sex, $F(3,2665) = 3.03$, $p = .048$, $\eta_p^2 = .002$. This was due to a significant main effect of Pain on RTs in males, $F(2,1413) = 8.14$, $p < .001$, $\eta_p^2 = .011$, but not in females, $F(2,1372) = 1.65$, $p = .193$, $\eta_p^2 = .002$. In males, participants with current pain had longer RTs (M = 1334, SD = 380) than those with no pain (M = 1279, SD = 390), $p = .015$, who in turn had longer RTs than those with recurrent pain (M = 1210, SD = 349), $p = .020$.

All other interactions were non-significant, ps > .095.

9. Study 3 discussion

A main effect or Trial Type suggested that there was an accuracy switch cost across all tasks. However, there was no evidence of an increase in the accuracy switch cost in the pain group on any of the tasks. On the AltBinary task, women without pain showed an accuracy switch cost while women with recurrent or current pain did not.

There was a significant RT switch cost across all tasks, and although participants in pain had longer RTs than those with recurrent pain or no pain (particularly so in males), again, pain did not increase switch costs.

10. General discussion

Consistent with theory [13] and numerous previous studies [21; 29; 31], we found that pain was associated with disrupted attention task performance, but that pain did not increase switch costs specifically. Rather, it had a small dampening effect on performance overall. In Study 1, laboratory induced thermal pain did not increase switch costs on two versions of an alternating runs switching paradigm and did not significantly affect accuracies or RTs overall. In Study 2, current pain in a large general population sample was associated with significantly lower accuracy and longer RTs on three versions of the switching task, one employing the task-cueing paradigm, and two employing the alternating runs paradigm either with or without cues, compared to participants with no pain. Interestingly, participants with recurrent pain conditions that were not causing pain at the time of the study did not significantly differ from either the current pain or no pain groups; their RTs fell in between those of the other two groups. This suggests that only current pain has negative effects on performance on these tasks. In Study 3, participants completed one of four versions of a more complex switching task, with either task-cueing or alternating runs switches, and with binary or numerical responses. Accuracy did not significantly differ between the three groups. However, current pain was associated with significantly longer RTs on all tasks

compared to the no pain and recurrent pain groups. In multiple tasks from Studies 2 and 3, accuracies decreased and RTs increased as pain intensity increased (see supplementary material). However, in all cases pain affected switch and repeat trials equally; we did not find an increase in switch costs on any of the tasks in any of the studies, contrary to van Ryckeghem et al [39].

Eccleston predicted that attention would be most susceptible to intrusion from pain when we switch our focus between tasks [12]. Our results have consistently failed to support this hypothesis, as have several previous studies [21; 29; 31]. However, van Ryckeghem et al. did find that pain increased switch costs on particular trial types. In their task-cueing paradigm, participants switched between three tasks, one of which was accompanied by a painful electrocutaneous stimulus on one quarter of trials. Participants were slower to switch away from a pain-related task onto a neutral task than they were to switch between two neutral tasks. This suggests that it is difficult to shift our attention away from pain, whereas our findings and those of previous studies have failed to show that it is more difficult to switch our attention between neutral tasks whilst we are in pain compared to when we are pain free (i.e. these studies address different questions).

Finding a dampening effect, as opposed to an increase in switch costs, suggests that pain interrupts attention consistently across the task, whilst participants' attention is engaged in a trial, not only when attention has been disengaged for shifting to a new task set. Given that humans can spend a significant amount of time mind-wandering, even from difficult or engaging tasks [22], perhaps it is not surprising that pain interrupts us on repeat trials as well as on switch trials. Kucyi, Salomons and Davis found that participants who reported frequent mind-wandering towards pain tended to show a greater increase in reaction times on a cognitive task during pain induction [23]. Perhaps, given that our minds frequently stray from tasks even when we are not in pain, the addition of pain simply gives a

compelling and difficult-to-disengage-from alternative topic for our minds to wander to, which it does consistently over the course of an attention task. Several large-scale studies have now shown that pain disrupts attention in a manner that dampens overall task performance ([3], Studies 2 and 3 in the current paper), although the effects tend to be small and may not be clinically relevant in isolation. Nevertheless, there could be important consequences in situations where small errors can lead to devastating consequences, such as driving or handling machinery.

Studies 2 and 3 also allowed us to examine the effects of analgesics on task performance (see supplementary material). This is important because if pain disrupts attention, then the next logical question is whether this disruption can be repaired by analgesics. In Study 2, participants who were in pain had slower RTs whether they were currently on analgesics or not. In Study 3, participants in pain who had taken analgesics had slower RTs than those who had not taken analgesics or those who were not in pain. While these results are not entirely consistent and participants were on a wide range of analgesics, both studies suggested that having taken analgesics did not restore response speed to the level of participants who were not in pain.

Consistent with previous findings [14; 18; 19; 32], females were more likely to report current and recurrent pain than males in our Internet samples from Studies 2 and 3, and women with recurrent pain also reported it to be of a higher intensity than males with recurrent pain. However, this did not translate into women with pain being more susceptible to attentional disruption than men with pain. In fact, in Study 3 the negative effect of pain on RTs could be isolated to men. We also found sex differences in strategies for coping with pain in Studies 2 and 3 (see supplementary material). In both studies, women were more likely to report using pain killers, hot/cold treatments, distraction, herbal remedies, relaxation/meditation, and going to bed as ways to deal with their pain. In both studies, men

were more likely than women to report using alcohol/drugs and in Study 3 they were more likely to say they would ignore their pain (see Keogh [20] for a discussion of men's pain coping behaviours). Overall, women reported greater use of coping strategies than men, which replicates previous findings [38]. Men possibly feel that they are able to cope with their pain without external help, or may be less likely to report their use of coping strategies.

There are several limitations to the studies reported here. Study 1 used laboratory-induced thermal pain in a small sample of university staff and students. The sample was therefore fairly homogenous, and the pain was low in threat value because participants were able to set the temperature to their threshold and to stop the pain at any time. We addressed these limitations in Studies 2 and 3 by using a large heterogeneous sample recruited via the Internet. This meant that the sample was more representative of the general population and that participants were experiencing naturally-occurring pain. However, data collection via the Internet brings its own set of issues. We cannot be sure that participants have responded honestly, although research suggests that most do [7; 25; 37], and we cannot control the environment in which participants completed the study, although our data suggested that only noise and effort affected task performance. However, the variety of settings in which participants completed the study could also be seen as a strength: the effect of pain on attention came through despite the lack of experimental control. Finally, given that we wanted to keep the online studies as short as possible, we did not ask participants about any other health factors that may be confounded with pain and may affect attention, such as sleep habits. Our view is that if an effect of pain on a given aspect of cognition can be identified with lab-induced pain in healthy volunteers and replicated with large heterogeneous samples, then at that stage it becomes fruitful to begin examining possible mediating variables in the heterogeneous samples.

An important direction for future research will be to examine the effects of pain on more ecologically valid attention tasks. Several studies have suggested that pain disrupts basic attention, but a remaining question is what happens when we complete complex tasks requiring multiple aspects of attention while experiencing pain. It may be the case that the effects of pain on the basic components of attention can be overcome, especially given how small the effects were in these studies (for example, the pain effect on accuracy did not reach significance in Study 3, and this may be because the additional task complexity led to greater engagement, which served as a distraction from pain). Alternatively, the effects may be cumulative leading to more severe disruption to everyday cognition.

In summary, we have found evidence that pain dampens accuracy and/or RTs on switching tasks overall, but no evidence that pain increases the cost of switching between tasks specifically. A picture is emerging from the literature in which pain has a small but significant negative effect on attention. This effect appears to be consistent across different types of pain. An important remaining question is the extent to which this impacts on daily life, not only through disrupted basic attentional processes but also through the impact this may have on higher level processes such as decision making and problem solving.

1 Acknowledgements

2 This work was supported by an unrestricted grant for research from Reckitt Benckiser

3 Healthcare (UK) Limited. The authors declare no conflicts of interest.

4

References

[1] Arnold LM, Crofford LJ, Mease PJ, Burgess SM, Palmer SC, Abetz L, Martin SA. Patient perspectives on the impact of fibromyalgia. Patient education and counseling 2008;73(1):114-120.

[2] Attridge N, Crombez G, Van Ryckeghem D, Keogh E, Eccleston C. The Experience of Cognitive Intrusion of Pain: scale development and validation. Pain 2015.

[3] Attridge N, Noonan D, Eccleston C, Keogh E. The disruptive effects of pain on n-back task performance in a large general population sample. Pain 2015.

[4] Aust F, Diedenhofen B, Ullrich S, Musch J. Seriousness checks are useful to improve data validity in online research. Behavior research methods 2013;45(2):527-535.

[5] Barnhoorn JS, Haasnoot E, Bocanegra BR, van Steenbergen H. QRTEngine: An easy solution for running online reaction time experiments using Qualtrics. Behav Res 2014:1-12.

[6] Bennett RM, Jones J, Turk DC, Russell IJ, Matallana L. An internet survey of 2,596 people with fibromyalgia. BMC musculoskeletal disorders 2007;8(1):27.

[7] Berinsky AJ, Margolis MF, Sances MW. Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self- Administered Surveys. American Journal of Political Science 2014;58(3):739-753.

[8] Bowering KJ, Butler DS, Fulton IJ, Moseley GL. Motor imagery in people with a history of back pain, current back pain, both, or neither. The Clinical journal of pain 2014;30(12):1070-1075.

[9] Buhrmester M, Kwang T, Gosling SD. Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? Perspectives on psychological science 2011;6(1):3-5.

[10] Clayton JA. Studying both sexes: a guiding principle for biomedicine. The FASEB Journal 2015:fj. 15-279554.

[11] Crombez G, Eccleston C, Baeyens F, Eelen P. Attentional disruption is enhanced by the threat of pain. Behav Res Ther 1998;36(2):195-204.

[12] Eccleston C. Chronic pain and distraction: an experimental investigation into the role of sustained and shifting attention in the processing of chronic persistent pain. Behaviour research and therapy 1995;33(4):391-405.

[13] Eccleston C, Crombez G. Pain demands attention: A cognitive–affective model of the interruptive function of pain. Psychological Bulletin 1999;125(3):356-366.

[14] Fillingim RB, King CD, Ribeiro-Dasilva MC, Rahim-Williams B, Riley Iii JL. Sex, Gender, and Pain: A Review of Recent Clinical and Experimental Findings. The Journal of Pain 2009;10(5):447-485.

[15] Greenspan JD, Craft RM, LeResche L, Arendt-Nielsen L, Berkley KJ, Fillingim RB, Gold MS, Holdcroft A, Lautenbacher S, Mayer EA. Studying sex and gender differences in pain and analgesia: a consensus report. Pain 2007;132:S26-S45.

[16] Ipeirotis PG. Demographics of mechanical turk. 2010.

[17] Katz RS, Heard AR, Mills M, Leavitt F. The prevalence and clinical impact of reported cognitive difficulties (fibrofog) in patients with rheumatic disease with and without fibromyalgia. JCR: Journal of Clinical Rheumatology 2004;10(2):53-58.

[18] Keogh E. Sex and gender differences in pain: a selective review of biological and psychosocial factors. The journal of men's health & gender 2006;3(3):236-243.

[19] Keogh E. Gender differences in the nonverbal communication of pain: A new direction for sex, gender, and pain research? Pain 2014;155(10):1927-1931.

[20] Keogh E. Men, masculinity, and pain. Pain 2015;156(12):2408-2412.

[21] Keogh E, Cavill R, Moore DJ, Eccleston C. The effects of menstrual-related pain on attentional interference. PAIN® 2014;155(4):821-827.

[22] Killingsworth MA, Gilbert DT. A wandering mind is an unhappy mind. Science 2010;330(6006):932-932.

[23] Kucyi A, Salomons TV, Davis KD. Mind wandering away from pain dynamically engages antinociceptive and default mode brain networks. Proceedings of the National Academy of Sciences 2013;110(46):18692-18697.

[24] Legrain V, Damme SV, Eccleston C, Davis KD, Seminowicz DA, Crombez G. A neurocognitive model of attention to pain: behavioral and neuroimaging evidence. Pain 2009;144(3):230-232.

[25] Mason W, Suri S. Conducting behavioral research on Amazon's Mechanical Turk. Behav Res 2012;44(1):1-23.

[26] Mason W, Watts DJ. Financial incentives and the performance of crowds. ACM SigKDD Explorations Newsletter 2010;11(2):100-108.

[27] Meiran N. Reconfiguration of processing mode prior to task performance. Journal of Experimental Psychology: Learning, Memory, and Cognition 1996;22(6):1423-1442.

[28] Monsell S. Task switching. Trends in cognitive sciences 2003;7(3):134-140.

[29] Moore DJ, Keogh E, Eccleston C. The interruptive effect of pain on attention. The Quarterly Journal of Experimental Psychology 2012;65(3):565-586.

[30] Moore DJ, Keogh E, Eccleston C. The effect of threat on attentional interruption by pain. Pain 2013;154(1):82-88.

[31] Moore DJ, Keogh E, Eccleston C. Headache impairs attentional performance. Pain 2013;154(9):1840-1845.

[32] Paller CJ, Campbell CM, Edwards RR, Dobs AS. Sex-Based Differences in Pain Perception and Treatment. Pain Medicine 2009;10(2):289-299.

[33] Paolacci G, Chandler J. Inside the turk understanding mechanical turk as a participant pool. Current Directions in Psychological Science 2014;23(3):184-188.

[34] Qualtrics L. Qualtrics [software], 2014.

[35] Rogers RD, Monsell S. Costs of a predictible switch between simple cognitive tasks. Journal of experimental psychology: General 1995;124(2):207.

[36] Schneider W, Eschman A, Zuccolotto A. E-prime User's Guide. Pittsburgh, USA: Psychology Software Tools Inc, 2012.

[37] Shapiro DN, Chandler J, Mueller PA. Using Mechanical Turk to study clinical populations. Clinical Psychological Science 2013:2167702612469015.

[38] Unruh AM, Ritchie J, Merskey H. Does gender affect appraisal of pain and pain coping strategies? The Clinical journal of pain 1999;15(1):31-40.

[39] Van Ryckeghem DM, Crombez G, Eccleston C, Liefooghe B, Van Damme S. The interruptive effect of pain in a multitask environment: an experimental investigation. The Journal of Pain 2012;13(2):131-138.

Figure 1a. Presentation sequence in the cued task in Study 1.

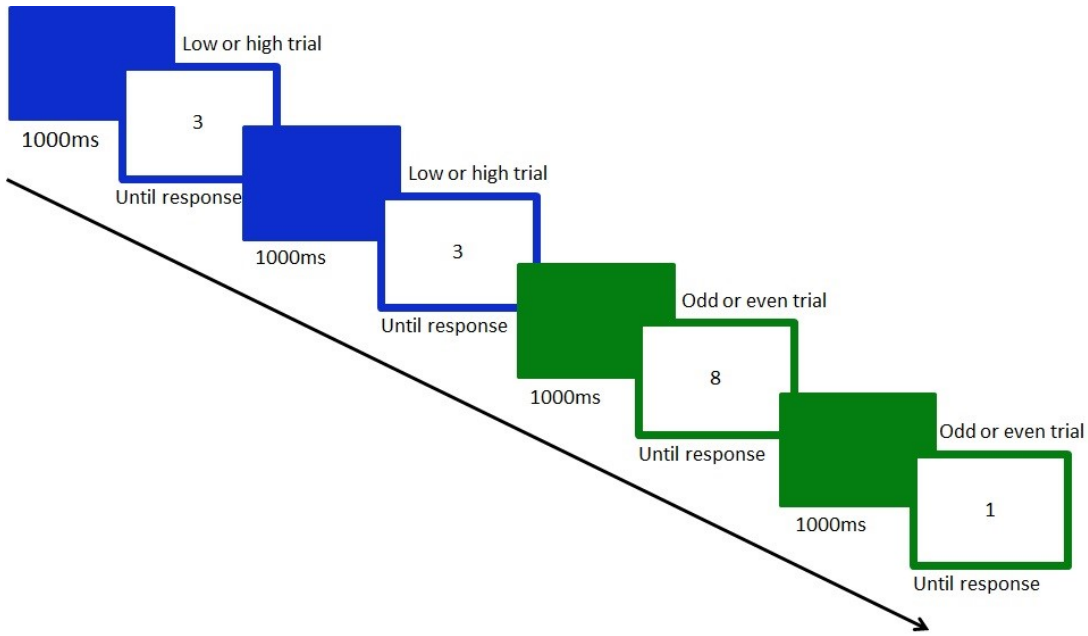

Figure 1b. Presentation sequence in the uncued task in Study 1.

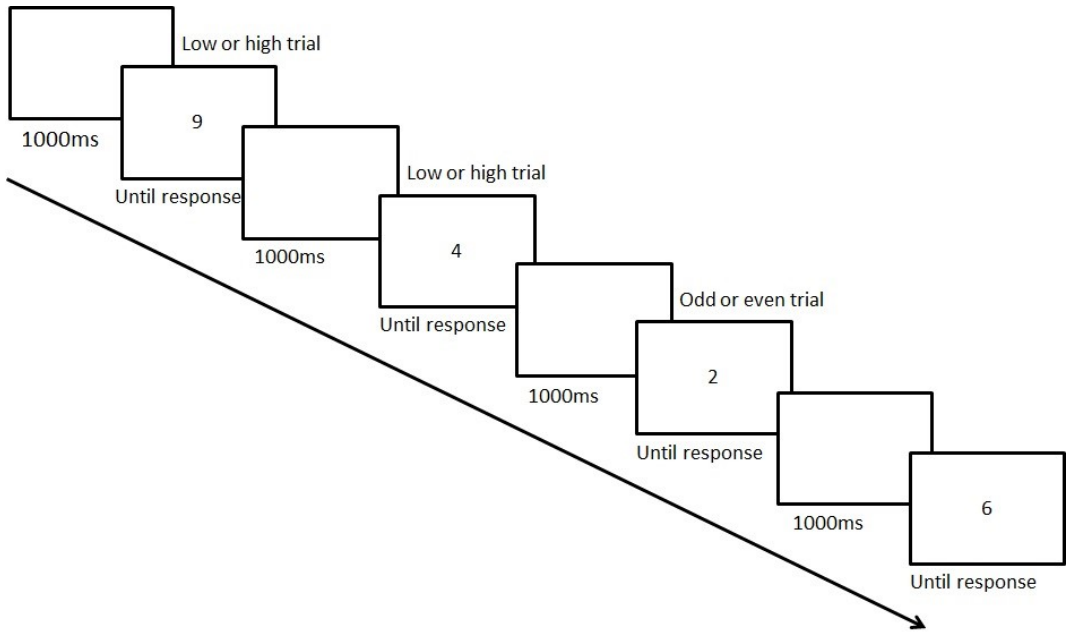

Figure 2. Example of the stimuli used in Study 3. Participants saw two cards and compared either the values on the cards or the number of digits on the cards.

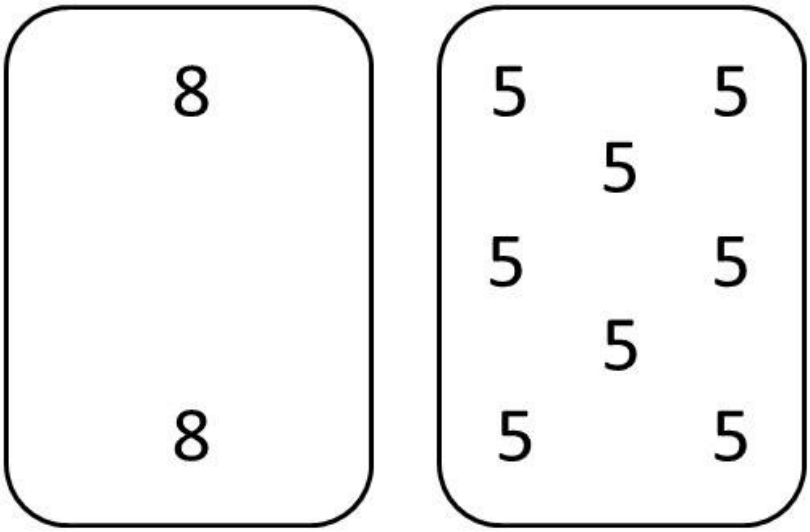

Table 1. Mean accuracy scores and reaction times (RTs) on the cued and uncued switching tasks in the pain and no pain conditions in Study 1 (standard deviations in parentheses).

		Cued		Uncued		Average
		Repeat	Switch	Repeat	Switch	
Accuracies	No pain	.977 (.03)	.972 (.02)	.939 (.09)	.925 (.10)	.953 (.043)
	Pain	.971 (.02)	.978 (.02)	.921 (.10)	.918 (.11)	.947 (.048)
	Pain cost	*-.003*	*+.006*	*-.018*	*-.017*	*-.006*
RTs (ms)	No pain	865 (261)	860 (262)	758 (221)	950 (236)	858 (211)
	Pain	852 (238)	934 (283)	770 (189)	966 (215)	881 (196)
	Pain cost	*-13*	*+74*	*+12*	*+16*	*+23*

Table 2. Number of participants in Study 2 who selected each type and duration of pain.

		Current pain (N)	Recurrent pain (N)
Duration of pain	Up to an hour	20	45
	Up to a day	24	34
	Up to a week	75	47
	Up to a month	24	32
	Up to three months	40	96
	Up to six months	56	147
	Up to a year	38	103
	Up to a decade	51	0
	Over a decade	83	0
Type of pain	Arthritis	60	72
	Upper back pain	71	111
	Lower back pain	72	71
	Non-muscular back pain	28	70
	Muscular non-back pain	26	28
	Hangover	9	17
	Headache	112	167
	Migraine	11	83
	Menstrual pain	19	73
	Nerve pain	27	32
	Joint pain	113	133
	Postsurgical	11	11
	Sciatica	21	44
	Stomach pain	25	64
	Sore throat	22	21
	Mouth/dental pain	38	68

Table 3. Mean accuracy scores and reaction times (RTs) in the current pain, recurrent but not current pain and no pain groups on the three switching tasks in Study 2 (standard deviations in parentheses).

		Unpredictable cued		Predictable cued		Predictable uncued		Average
		Repeat	Switch	Repeat	Switch	Repeat	Switch	
Proportion of trials correct	Current Pain	.943 (.063)	.911 (.078)	.937 (.078)	.908 (.078)	.817 (.141)	.816 (.125)	.889 (.063)
	Recurrent Pain	.956 (.066)	.931 (.075)	.958 (.075)	.932 (.075)	.853 (.132)	.845 (.123)	.912 (.066)
	No pain	.952 (.066)	.920 (.083)	.947 (.066)	.913 (.083)	.834 (.133)	.842 (.117)	.901 (.066)
RTs in milliseconds	Current Pain	908 (215)	964 (231)	904 (214)	1017 (241)	933 (231)	1080 (278)	968 (197)
	Recurrent Pain	885 (214)	956 (230)	871 (213)	994 (240)	920 (231)	1059 (277)	948 (196)
	No pain	845 (216)	899 (232)	853 (216)	951 (243)	887 (233)	1009 (280)	908 (198)

Table 4. Number of participants in Study 3 who selected each type and duration of pain.

		Current pain (N)	Recurrent pain (N)
Duration of pain	Up to an hour	133	0
	Up to a day	236	0
	Up to a week	200	108
	Up to a month	85	98
	Up to three months	84	109
	Up to six months	58	101
	Up to a year	340	335
	Up to a decade	0	489
	Over a decade	340	297
Type of pain	Arthritis	161	183
	Upper back pain	210	239
	Lower back pain	195	275
	Non-muscular back pain	348	401
	Muscular non-back pain	90	108
	Hangover	30	60
	Headache	332	491
	Migraine	46	236
	Menstrual pain	69	264
	Nerve pain	81	110
	Joint pain	366	390
	Postsurgical	29	47
	Sciatica	56	92
	Stomach pain	122	212
	Sore throat	38	57
	Mouth/dental pain	139	242

Table 5. Mean accuracy scores and reaction times (RTs) in the current pain, recurrent but not current pain and no pain groups on the four switching tasks in Study 3 (standard deviations in parentheses).

		AltBinary		AltNum		RandBinary		RandNum		Average
		Repeat	Switch	Repeat	Switch	Repeat	Switch	Repeat	Switch	
Proportion of trials correct	Current pain	.867 (.199)	.861 (.184)	.566 (.196)	.553 (.196)	.912 (.196)	.906 (.196)	.622 (.195)	.613 (.195)	.737 (.197)
	Recurrent pain	.863 (.199)	.855 (.189)	.544 (.191)	.538 (.191)	.926 (.191)	.916 (.191)	.677 (.200)	.666 (.189)	.748 (.197)
	No pain	.857 (.193)	.863 (.193)	.575 (.193)	.558 (.193)	.927 (.202)	.917 (.184)	.683 (.202)	.670 (.202)	.756 (.180)
RTs in milliseconds	Current pain	1098 (275)	1192 (287)	1570 (275)	1667 (287)	1050 (277)	1074 (289)	1512 (275)	1553 (287)	1340 (279)
	Recurrent pain	1050 (274)	1134 (286)	1531 (277)	1627 (289)	979 (276)	1006 (288)	1502 (277)	1542 (289)	1296 (280)
	No pain	1061 (279)	1132 (292)	1571 (276)	1665 (288)	953 (276)	979 (288)	1487 (277)	1538 (289)	1298 (279)

Supplementary material

1. Study 2 additional analyses

1.1. Coping strategies

Participants indicated how they usually dealt with pain by ticking all applicable strategies from a list of options or entering their own response into a free text box. Sex differences in the use of each coping strategy were investigated with a series of Chi Square tests. The most common response was pain killers ($N = 588$), which were used by a higher percentage of females (58.8%) than males (50.7%), $\chi^2 = 7.00$, $p = .008$. The second most common response was ignoring the pain ($N = 514$), which did not differ in prevalence between males (50.5%) and females (44.9%), $\chi^2 = 3.43$, $p = .064$. Going to bed was the next most common selection ($N = 424$), and was more common in females (43.2%) than in males (35.7%), $\chi^2 = 6.18$, $p = .013$. Hot/cold treatments were the next most popular option ($N = 356$), and were used by more females (40.1%) than males (26.5%), $\chi^2 = 22.20$, $p < .001$. Relaxation/meditation ($N = 332$) was more common in females (34.5%) than in males (27.4%), $\chi^2 = 6.25$, $p = .012$. Exercise ($N = 324$), was also more common in females (33.1%) than in males (27.3%), $\chi^2 = 4.41$, $p = .036$. Distraction ($N = 302$) was also more common in females (31.6%) than in males (25.3%), $\chi^2 = 5.29$, $p = .022$. Less popular options were herbal remedies ($N = 137$), which were more common in females (14.6%) than in males (10.5%), $\chi^2 = 4.27$, $p = .039$, using alcohol/drugs ($N = 69$), which was more common in males (8.3%) than in females (4.2%), $\chi^2 = 7.46$ $p = .006$, and acupuncture ($N = 22$), which did not differ in prevalence between the sexes (female prevalence = 2.1%; male prevalence = 2.0%), $\chi^2 = .02$, $p = .877$. Other strategies reported include yoga, stretches, pressure/massage, praying, animal therapy and avoiding movement.

Participants were also asked to indicate how often they generally take pain killers by selecting one of the following options: 1 = every day (N = 84), 2 = a few times per week (N = 157), 3 = once per week (N = 66), 4 = a few times monthly (N = 169), 5 = less than monthly (N = 269) or 6 = never (N = 308). A Mann-Whitney U test showed that females reported taking pain killers more frequently (median = 4, i.e. a few times monthly) than males (median = 5, i.e. less than monthly), U = 116.699.5, z = -4.32, p < .001.

1.2. Effects of pain intensity, location and duration on task performance

The relationship between current pain intensity and task performance was investigated by correlating pain intensity with accuracy and RTs on each of the three tasks. For the cued unpredictable task, pain intensity was significantly negatively correlated with accuracies, $r(384) = -.22$, $p < .001$, and significantly positively correlated with RTs, $r(377) = .19$, $p < .001$. For the cued predictable task, pain intensity was again significantly negatively correlated with accuracies, $r(384) = -.27$, $p < .001$, and positively correlated with RTs, $r(374) = .20$, $p < .001$. For the uncued predictable task, pain intensity was not significantly correlated with accuracies, $r(378) = .09$, $p = .076$, or RTs, $r(258) = .03$, $p = .687$.

Participants indicated which type(s) of pain they were experiencing by checking boxes from a list of options, or by entering their own response in a free text box. Participants were then organised into 'pain type groups': arthritis only (N = 9), upper back/neck pain only (N = 18), lower back pain only (N = 17), non-muscular back pain only (N = 32), muscular non-back pain only (N = 9), hangover only (N = 1), headache only (N = 27), migraine only (N = 1), menstrual pain only (N = 6), nerve pain only (N = 2), joint pain only (N = 31), post-surgical pain only (N = 2), stomach pain only (N = 8), throat pain only (N = 6), tooth and dental pain only (N = 14) or multiple pain types (N = 265).

We investigated task performance in the pain type groups that contained at least 10 participants with data on all three tasks (non-muscular back pain, $N = 18$, headache, $N = 16$, joint pain, $N = 14$ and multiple pain types, $N = 173$) using 4 (pain group) × 3 (task: cued unpredictable, cued predictable, uncued predictable) × 2 (trial type: switch, repeat) ANOVA for accuracy and RT scores (sex was not included due to small cell counts). For accuracies, we found a significant main effect of pain group, $F(3,217) = 4.02$, $p = .008$, $\eta_p^2 = .053$, and post hoc LSD tests showed that the joint pain group scored significantly lower ($M = 8.28$, $SD = .106$) than the multiple pains group ($M = .892$, $SD = .081$), $p = .001$. The non-muscular back pain ($M = .875$, $SE = .096$) and headache ($M = .867$, $SE = .094$) groups did not differ from any other groups, ps > .062. There were no interactions between pain group and task or trial type. For RTs, there was no main effect of pain group and no interactions.

Next we examined the relationship between current pain duration and task performance using Spearman's rank correlations. Duration was not correlated with accuracies or RTs on any of the tasks, all rs < ±.09, all ps > .086.

1.3. Effects of analgesics on task performance

Participants who reported being in pain at the time of the study had lower accuracy scores and longer RTs than those who were not in pain, and on two of the three tasks higher intensity pain was associated with worse task performance. Next we investigated the effects of analgesics on accuracies and RTs using 3 (group: no pain and no analgesics, pain and no analgesics, and pain and analgesics) × 3 (task: cued unpredictable, cued predictable, uncued predictable) × 2 (trial type: switch, repeat) × 2 (Sex) mixed ANOVAs. Participants reported taking a range of analgesics, including but not limited to: aspirin, codeine, ibuprofen, methodone, naproxen, paracetamol and tramadol. Splitting the participants by analgesic type

would lead to small cell sizes, so here we simply consider whether or not they reported having taken analgesics at the time of the study.

In the full sample, 632 participants reported no pain and no analgesics, 32 reported no pain and analgesics (excluded here to allow sufficient cell sizes to examine sex differences), 275 reported pain and no analgesics, and 139 reported pain and analgesics. However, for these analyses only participants who completed all three tasks were included and the group sizes were 347, 162 and 84, respectively. Because the Task and Trial Type effects were reported above, here we only report effects of analgesic group.

For accuracy scores, there was a significant main effect of analgesic group, $F(2,587) = 4.98$, $p = .007$, $\eta_p^2 = .017$. Post hoc LSD tests showed that participants who reported no pain and no analgesics were more accurate ($M = .905$, $SD = .750$) than participants who reported pain and had taken analgesics ($M = .882$, $SD = .640$), $p = .006$, or reported pain and no analgesics ($M = .892$, $SD = .640$), $p = .033$. There was no difference between participants with pain who had or had not taken analgesics, $p = .306$. There was also a three way interaction between Task, Sex and Analgesics Group, $F(4,1174) = 2.59$, $p = .036$, $\eta_p^2 = .009$. This was due to a borderline-significant interaction between Sex and Analgesic group on the uncued task, $F(2,626) = 2.99$, $p = .051$, and no interaction on the random, $F(2,937) = .297$, $p = .743$, or cued tasks, $F(2,935) = .35$, $p = .702$. The borderline interaction on the uncued task was due to Analgesics group predicting task performance in males, $F(2,328) = 4.46$, $p = .012$, but not in females, $F(2,298) = 2.30$, $p = .102$. Bonferroni post hoc tests showed that in males, participants who had pain and had not taken analgesics scored significantly lower ($M = .793$, $SD = .121$) than participants who did not have pain and had not taken analgesics ($M = .834$, $SD = .122$), $p = .028$. There were no other significant group comparisons. There were no interactions between analgesic group and any other factor.

For RTs, there was again a significant main effect of analgesic group, $F(3,587) = 6.35$, $p = .002$, $\eta_p^2 = .021$. Post hoc LSD tests showed the same pattern of effects as with the accuracy scores: the pain and analgesics group (M = 990ms, SD = 205ms), and the pain and no analgesics group (M = 966ms, SD = 199ms) had significantly longer RTs than the no pain no analgesics group (M = 916ms, SD = 201ms), $p = .001$ and $p = .006$, respectively. There were no interactions between analgesic group and any other factor.

1.4. Study 2 supplementary analyses discussion

Higher pain intensity was related to worse overall performance on the task cueing and cued alternating runs tasks, but duration of pain was not related to performance. This pattern of results was also found by Attridge et al [3] in their investigation of the effect of pain on n-back task performance in a large internet sample.

Our large heterogeneous sample allowed us to investigate the relationship between analgesics and attentional disruption for the first time. Participants who reported no pain and no analgesics scored significantly higher than participants who reported pain and had taken analgesics, or reported pain and no analgesics. For RTs, the pain and analgesics group and the pain and no analgesics group had significantly longer RTs than the no pain no analgesics group. Overall this suggests that pain is associated with worse task performance than no pain, regardless of whether the participant had taken analgesics. We therefore found no evidence that analgesics further disrupt nor repair attentional disruption from pain.

2. Study 3 additional analyses

2.1. Coping strategies

Participants were indicated how they usually dealt with pain by ticking a list of options or entering their own response into a free text box. The most common response was

painkillers (N = 1722), which were used by a higher percentage of women (59.8%) than men (49.8%), χ^2 = 31.34, p < .001. This was followed by ignoring the pain (N = 1551), which was more common in men (51.5%) than in women (47%), χ^2 = 6.58, p = .010. Next most common was going to bed (N = 1325), which was selected by more women (45.8%) than men (38.6%), χ^2 = 16.90, p < .001. Hot/cold treatments were next most popular (N = 1000), and were more common in women (38.9%) than in men (25%), χ^2 = 70.16, p < .001. Relaxation/meditation (N = 996), was also selected by a higher proportion of women (33.6%) than men (29.8%), χ^2 = 5.41, p = .020. Exercise (N = 936) did not differ in prevalence between women (29.5%) and men (30.0%), χ^2 = .11, p = .742. Distraction (N = 894) was more prevalent in women (33%) than in men (24%), χ^2 = 30.97, p < .001, as were herbal remedies (N = 398, 14.5% of women, 10.9% of men), χ^2 = 9.50, p = .002. Alcohol/drugs (N = 240) was the only coping strategy that was more popular with men (10%) than with women (5.1%), χ^2 = 27.10, p < .001. Finally, acupuncture (N = 79) did not differ in prevalence between women (2.3%) and men (2.7%), χ^2 = .40, p = .529.

Participants were also asked to indicate how often they generally take painkillers from the following options: 1 = every day (N = 263), 2 = a few times per week (N = 480), 3 = once per week (N = 213), 4 = a few times monthly (N = 447), 5 = less than monthly (N = 861) or 6 = never (N = 821). A Mann-Whitney U test showed that females reported taking pain killers more frequently (median = 4, i.e. a few times monthly) than males (median = 5, i.e. less than monthly), U = 1,033,453.5, z = -8.20, p < .001.

2.2. Effects of pain intensity, location and duration on task performance

The relationship between current pain intensity and task performance was investigated by correlating pain intensity with accuracy and RTs on each of the four tasks separately, since the tasks varied so widely in difficulty. For the AltBinary task, pain intensity was

significantly negatively correlated with accuracy, $r(251) = -.29$, $p < .001$, and not correlated with RTs, $r(250) = .10$, $p = .107$. For the RandBinary task, pain intensity was again significantly negatively correlated with accuracies, $r(335) = -.27$, $p < .001$, and positively correlated with RTs, $r(334) = .14$, $p = .008$. For the AltNum task, pain intensity was significantly negatively correlated with accuracies, $r(311) = -.29$, $p < .001$, but was not correlated with RTs, $r(249) = -.004$, $p = .952$. For the RandNum task, pain intensity was significantly negatively correlated with accuracy, $r(297) = -.17$, $p = .004$, and significantly positively correlated with RTs, $r(252) = .16$, $p = .011$.

Participants indicated which type(s) of pain they were experiencing by checking boxes from a list of options, or by entering their own response in a free text box. Participants were then organised into 'pain type groups': arthritis only ($N = 13$), upper back/neck pain only ($N = 43$), lower back pain only ($N = 43$), non-muscular back pain only ($N = 98$), muscular non-back pain only ($N = 24$), hangover only ($N = 2$), headache only ($N = 93$), migraine only ($N = 8$), menstrual pain only ($N = 8$), nerve pain only ($N = 14$), joint pain only ($N = 85$), post-surgical pain only ($N = 7$), stomach pain only ($N = 29$), throat pain only ($N = 4$), tooth and dental pain only ($N = 41$) or multiple pain types ($N = 625$). We investigated accuracies and RTs on each task in the pain type groups that contained at least 10 participants for that task.

For the AltBinary task we ran two 5 (pain group: non-muscular back, $N = 15$; headache, $N = 18$; joint pain, $N = 21$, lower back muscular, $N = 13$; upper back muscular, $N = 13$) x 2 (trial type) ANOVAs, one on accuracies and one on RTs. For accuracies, there was no main effect of trial type, $F(1,75) = 1.32$, $p = .254$, $\eta_p^2 = .017$, no main effect of pain type, $F(4,75) = 1.37$, $p = .253$, $\eta_p^2 = .068$, and no interaction, $F(4,75) = .18$, $p = .948$, $\eta_p^2 = .010$. For RTs, there was a significant main effect of trial type, $F(1,75) = 34.87$, $p < .001$, $\eta_p^2 = .317$, but no main effect of pain type, $F(4,75) = 1.00$, $p = .412$, $\eta_p^2 = .051$, and no interaction, $F(4,75) = 1.44$, $p = .230$, $\eta_p^2 = .071$.

For the RandBinary task we ran two 4 (pain group: non-muscular back, $N = 34$; headache, $N = 26$; joint pain, $N = 24$, upper back muscular, $N = 10$) x 2 (trial type) ANOVAs, one on accuracies and one on RTs. For accuracies, there was no main effect of trial type, $F(1,88) = .06$, $p = .807$, $\eta_p^2 = .001$, no main effect of pain type, $F(3,88) = 1.62$, $p = .191$, $\eta_p^2 = .052$, and no interaction, $F(3,88) = 2.21$, $p = .093$, $\eta_p^2 = .070$. For RTs, there was a main effect of trial type, $F(1,89) = 9.37$, $p = .003$, $\eta_p^2 = .095$, no main effect of pain type, $F(3,89) = .50$, $p = .685$, $\eta_p^2 = .017$.

For the AltNum task we ran two 6 (pain group: non-muscular back, $N = 29$; headache, $N = 26$; joint pain, $N = 20$, lower back muscular, $N = 11$; upper back muscular, $N = 10$; tooth/dental, $N = 14$) x 2 (trial type) ANOVAs, one on accuracies and one on RTs. For accuracies, there was no main effect of trial type, $F(1,97) = .93$, $p = .337$, $\eta_p^2 = .009$, no main effect of pain type, $F(5,97) = .64$, $p = .673$, $\eta_p^2 = .032$, and no interaction, $F(5,97) = .09$, $p = .995$, $\eta_p^2 = .004$. For RTs, there was a significant main effect of trial type, $F(1,78) = 39.35$, $p < .001$, $\eta_p^2 = .335$, but no main effect of pain type, $F(5,78) = 1.29$, $p = .276$, $\eta_p^2 = .076$, and no interaction, $F(5,78) = .31$, $p = .906$, $\eta_p^2 = .019$.

For the RandNum task we ran two 6 (pain group: non-muscular back, $N = 20$; headache, $N = 23$; joint pain, $N = 20$, lower back muscular, $N = 11$; upper back muscular, $N = 10$, stomach, $N = 10$) x 2 (trial type) ANOVAs, one on accuracies and one on RTs. For accuracies, there was a main effect of trial type, $F(1,84) = 7.58$, $p = .007$, $\eta_p^2 = .083$, but no main effect of pain type, $F(5,84) = .92$, $p = .470$, $\eta_p^2 = .052$, and no interaction, $F(5,84) = .82$, $p = .538$, $\eta_p^2 = .047$. For RTs, there was a significant main effect of trial type, $F(1,73) = 5.65$, $p = .020$, $\eta_p^2 = .072$, but no main effect of pain type, $F(5,73) = .91$, $p = .477$, $\eta_p^2 = .059$, and no interaction, $F(5,73) = .30$, $p = .911$, $\eta_p^2 = .020$.

Next we examined the relationship between current pain duration and task performance using Spearman's rank correlations. Duration was positively correlated with

RTs on the AltBinary, $r(670) = .10$, $p = .009$, and RandBinary tasks, $r(855) = .150$, $p < .001$. Duration was negatively correlated with accuracies on the RandBinary, $r(857) = -.10$, $p = .004$, and RandNum tasks, $r(816) = -.07$, $p = .044$. All other correlations were non-significant, rs < ±.038, ps > .325.

2.3. Effects of analgesics on task performance

Next we investigated the effects of analgesics on accuracies and RTs using 3 (analgesics group: no pain and no analgesics, $N = 1524$, pain and no analgesics, $N = 682$, and pain and analgesics, $N = 410$) × 4 (Task: AltBinary, AltNum, RandBinary, RandNum) × 2 (trial type: switch, repeat) × 2 (Sex) mixed ANOVAs. As in Study 2, participants reported taking a range of analgesics, including but not limited to: aspirin, codeine, ibuprofen, methodone, naproxen, paracetamol and tramadol. Splitting the participants by analgesic type would lead to small cell sizes, so here we simply consider whether or not they reported having taken analgesics at the time of the study. Because the Task and Trial Type effects were reported above, here we only report effects involving analgesic group.

For accuracy scores, there was no main effect of analgesic group, $F(2,2540) = 2.32$, $p = .098$, $\eta_p^2 = .002$ and no interactions between analgesic group and any other factor. For RTs, there was a significant main effect of analgesic group, $F(2,2298) = 10.17$, $p < .001$, $\eta_p^2 = .009$. Post hoc LSD tests showed that the pain and analgesics group ($M = 1374$ms, $SD = 285$) had significantly longer RTs than the pain and no analgesics group ($M = 1325$ms, $SD = 276$), $p = .010$, and the no pain no analgesics group ($M = 1299$ms, $SD = 277$), $p < .001$. No other comparisons reached significance. There were no interactions between analgesic group and any other factor.

2.4. Study 3 supplementary analyses discussion

The intensity of pain that participants reported predicted task performance. Higher intensity pain was associated with lower accuracies on all four tasks, and it was associated with longer RTs on the random switch tasks. The duration of participants' pain was also important. The longer participants had had their pain, the lower their accuracies on the random switch tasks, and the slower their RTs were on the binary response tasks. Finally, participants who had pain and had taken analgesics had significantly longer RTs than those who had pain and had not taken analgesics or those had no pain and had not taken analgesics.